

NEW EVIDENCE SUPPORTING THE ESP INTERPRETATION OF THE PRATT-WOODRUFF EXPERIMENT

By J. G. PRATT

ABSTRACT: Heretofore the criticism of the Pratt-Woodruff ESP experiment has been based on the discovery that the significant results were concentrated in those trials made in relation to the key cards that were in the end positions in the preceding run (E-cards). The critics attributed this finding to fraud on the part of one of the experimenters on the ground that he must have kept track of those key cards from one run to the next and misplaced cards to those positions to score spurious hits. The experimenters, on the other hand, argued that the E-cards could have been more favorable for ESP success because of the subject's psychological reaction to them. This paper describes an analysis based upon the E-card trials alone for the highest scoring subject, P.M. It was found that scoring was near the chance level for those E-cards that were not shifted from the positions they occupied in the preceding run and the results were significant only on the shifted E-cards. This finding supports the ESP interpretation of the E-card effect and it is not consistent with the Hansel interpretation of experimenter fraud. Other analyses also support the ESP hypothesis as the explanation of the results of the Pratt-Woodruff experiment.

This paper is a research report. Readers of the *Journal* will, I think, be relieved to know that they are not confronted with merely a renewal of the controversy that was first waged in these pages fifteen years ago (Hansel, 1961; Pratt & Woodruff, 1961) and reopened thirteen years later (Medhurst & Scott, 1974; Pratt, 1974a, 1974b; Scott, 1974) without resolving the basic issue (Gatlin, 1975). The present paper presents a new type of analysis of the data of P.M., the highest scoring subject in the experiment at issue (Pratt & Woodruff, 1939). In the analysis presented in this report one result was predicted if the special effect in the data discovered by Hansel (described in the next two sections) was due to the subject's ESP performance (ESP) and another result if it came about, as Hansel assumed, through card misplacement by Woodruff (CM). The present research is thus closely related to the controversy, and it is essential

I wish to express my sincere appreciation to the following persons who read an earlier draft of this paper with special attention to the validity of the statistical design and the interpretation given to the findings: Dr. J. A. Greenwood, Dr. T. N. E. Greville, Dr. John Palmer, Dr. Lila L. Gatlin, and Dr. Edward F. Kelly. I appreciate particularly their unanimous opinion that the study should be published, but I alone am responsible for any weaknesses that may be found in the final report.

that the reader should have in mind a few special definitions as well as the main features of the controversy itself. Readers who already have this information may want to turn immediately to the section with the subtitle "Description of Analysis and Predictions."

DEFINITIONS

STM procedure, Pratt-Woodruff style: A vertical screen placed on a small table divided it into two approximately equal areas. On one side sat the subject facing a permutation of the five ESP symbols, the key cards, that hung on a row of pegs across the face of the screen. In front of a narrow aperture across the bottom of the screen there was a row of five blank "marker" cards corresponding to the positions of the key cards. One experimenter, Woodruff, sat on the other side of the screen holding a shuffled pack of ESP cards (target cards) and watching the row of blank cards for the subject's pointer. In preparation for a new run, I, as second experimenter, who worked on the subject's side of the screen, removed the key cards from their pegs and handed them to the subject to be returned to the pegs in an order that was not known to Woodruff. When the key cards were in position and when Woodruff announced that he was ready, the subject pointed to the marker cards, one after another, to indicate where Woodruff should place the target cards. The subject's purpose was to have as many as possible of the target cards match the key cards. At the end of the run I recorded the order of the key cards while Woodruff recorded the ESP cards in the five piles back of the screen, and these two independent records were clipped together and deposited through a slot into a locked box. Then the screen was laid on its side and I, observed by Woodruff and the subject, checked the number of hits, i.e., the cards in the five piles that matched the key cards. Each experimenter independently recorded the total number of hits for the run. Then the screen was returned to its vertical position and the procedure was repeated for the next run.

E-cards: The two key cards in a given run which had occupied the two end positions in the row of pegs during the previous run.

E-piles: The cards laid down during the run by Woodruff back of the screen in the two positions corresponding to the key cards that were on the two end pegs during the preceding run.

M-cards: The three key cards in a given run which occupied the three central positions during the preceding run.

M-piles: The target cards placed by Woodruff back of the screen

in the three positions corresponding to the key cards that were on the three inner pegs during the preceding run.

Hansel effect: The tendency for the hits in the data of P.M., the highest-scoring subject in the Pratt-Woodruff experiment, to be concentrated in the E-piles.

The card misplacement (CM) hypothesis: Hansel's hypothesis that the effect he discovered was produced through Woodruff's keeping track of one or both of the E-cards and occasionally during the run identifying a target card that matched an E-card and placing it in that E-pile when the subject was pointing to one of the other positions.

Note: To avoid confusion in reading later parts of this paper, one needs to keep in mind the fact that the letters "E" and "M" always refer to the positions of key cards during the run immediately preceding the one in which the key cards and target cards are compared to obtain the run score. Heretofore, the controversy has been concerned with the difference in scoring rate between E-piles and M-piles. In this paper attention will be focused upon the E-piles alone. The level of scoring will be compared in those E-piles when the key cards, considered individually, were replaced in the same end positions as they had been in before (unshifted E-cards) and in the remaining E-piles when the key cards were shifted to one of the other four pegs (shifted E-cards).

SYNOPSIS OF THE CONTROVERSY

The rationale of the criticism that initiated the controversy was the following statement by Hansel (1961): "When examining an experiment of this nature in order to see whether it is foolproof, we first assume that ESP is impossible and then seek some other cause of the high scores" (p. 102). His search led him to the discovery that in the records of the highest-scoring subject, P.M., the significant results were largely concentrated in the E-piles. He interpreted this to mean that Woodruff had produced spurious hits in the E-piles through target card misplacement, as defined above. Hansel acknowledged that this would not have been possible if I had mixed the target cards after removing them from the pegs before handing them to the subject. I stated that I remembered changing the order of the key cards, but this point was not covered explicitly in our original report (Pratt & Woodruff, 1939). Hansel was therefore able to argue that I must have forgotten the fact that I did not mix up the key cards.

Pratt and Woodruff (1961) acknowledged that the E-pile effect was highly significant, but we argued that, because of the test conditions, it could more reasonably be interpreted as an ESP salience effect. We suggested that the subject's assigned task of putting the cards back on the pegs in a different order caused her to pay special attention to the cards that had been in the more prominent end positions, the E-cards. This special focusing of attention upon the end cards of the previous run made those symbols more salient as ESP targets and accounted for her higher rate of success on them. We showed that Hansel's claim that the E-pile effect was also significantly present in the combined data of the other four high-scoring subjects was not sound since there were flaws in his statistical analysis of those records.

Medhurst and Scott (1974) applied another statistical test to the records of the other high-scoring subjects and found that the E-pile scoring rate was higher than the M-pile rate in their records at an acceptable level of significance ($p = .012$, one-tailed) though at a much lower level of significance than this analysis gave for P.M. ($p < 2 \times 10^{-6}$, one-tailed). Medhurst and Scott concluded: "The evidence is not, of course, compelling. It is open to anyone to prefer the hypothesis that an unlikely coincidence has occurred or that the psychological peculiarity attributed by Pratt and Woodruff to the subject P.M. applied to more than one subject, or to produce yet another hypothesis in terms of an ESP effect. . . . However, it seems clear that the new evidence in this paper moves the balance at least some distance toward Hansel's hypothesis" (p. 182).

I replied (Pratt, 1974a) that the extension of the Hansel finding to the other high-scoring subjects did not shift the balance toward his hypothesis and that what was needed was some analysis that would discriminate between ESP and CM. I suggested one kind of correlation analysis between E- and M-pile scores within the run for P.M. that might serve this purpose, but did so without the opportunity to test the matter since the data had not been available to me after my departure from the Duke Parapsychology Laboratory in 1963.

Scott (1974) worked out my suggested correlation and found that it was significantly positive, as I had suggested it should be on the ESP hypothesis. However, he raised some logical objections which cast doubt upon whether this analysis would effectively discriminate between ESP and CM.

I replied (Pratt, 1974b) with the suggestion that perhaps a com-

puter study could be made to show whether the analysis did or did not effectively discriminate between the two hypotheses.

Gatlin (1975) carried out a computer simulation of the Pratt-Woodruff tests of P.M. with a Monte Carlo experiment and concluded that my proposed analysis was not an effective discriminator. She reached the conclusion, however, that both hypotheses, ESP and CM, were still viable. In referring to the ESP explanation of the E-card effect that Woodruff and I had proposed, Gatlin emphasized the process of short-term memory on the part of the subject in deciding upon the new order of the key cards. Her mentioning the possible role of memory sparked an idea in my mind which led me to a new kind of analysis to discriminate between ESP and CM.

DESCRIPTION OF ANALYSIS AND PREDICTIONS

This analysis sought to discriminate between ESP and CM by using a basis of differentiation that was confined to the E-cards and the E-piles. Because the E-card effect was so much stronger in P.M., I limited my research to her E-card trials. If evidence was found from her records that clearly pointed to an ESP interpretation of the effect discovered by Hansel, we would be safe in assuming that the same will be true for the much weaker E-card effect that might be present in the records of the other four high-scoring subjects as a group.

In the testing procedure of the Pratt-Woodruff experiment, the subject was assigned the task of replacing the key cards on the pegs before each run, and she was told to put them up in a different order from the one they had been in before. Obviously, if the subject took a naïve and literal view of the requirement to make the new order different from the old one, she would have been able to carry out the instructions to change the order only to the degree that she remembered what that previous order had been. The fact that the Hansel effect occurred can be interpreted as showing that P.M. primarily took account of the E-cards when she was rearranging the key cards on the pegs.

The new idea that occurred to me while reading the Gatlin paper was that the subject would not have remembered all of the key cards that were formerly in the end positions, and that when she failed to remember E-cards they should not have been different for her, psychologically, from M-cards. Those forgotten E-cards and all of the M-cards would have been replaced "randomly" upon the pegs

that were still empty after the remembered E-cards had been put up in their changed positions. Then it occurred to me that we should be able to identify at least *some* of the E-cards that she did not remember: those that were left for the new run in the same position as they had been in during the last run. Since she was making a special effort to change the positions of the key cards, she would have replaced an E-card on the same peg as before only when she did not remember that it had been there.

I therefore planned an analysis which would involve comparing the rates of scoring on E-cards which were shifted to a different position and on E-cards which were not shifted.

On the ESP hypothesis, unshifted E-cards, being psychologically like the M-cards, should also show a low rate of success. The shifted E-cards, on the other hand, should contain those cards that the subject consciously remembered and replaced on other pegs in carrying out the instructions given to her (although not all of the shifted E-cards would have been remembered; more about this point later). On the ESP hypothesis, I therefore predicted higher scoring on shifted E-cards than on unshifted ones.

On the CM hypothesis, on the other hand, there was no obvious reason to expect any difference in the rate of scoring on shifted and unshifted E-cards, since this distinction would have no meaning for Woodruff. If pressed to indicate which would give the better scoring rate through CM if any difference should be found, I would say that we should expect more success on the unshifted E-cards. The reason is that it might be easier for the experimenter to keep track of an E-card when it was returned to the same peg than when it was placed on a different one.

The statistical test that I decided to apply to the data with the E-card trials divided into these two categories is a chi-square test based upon a two-by-two contingency table, the same evaluation as was used by Medhurst and Scott.

The research plans outlined above were made to the point that has now been reached in this report before the data were examined. Thus the analysis was worked out along logical and objective lines rather than being arrived at "after the fact" on the basis of endless poring over the records in search of some trend in the data that could be construed as supporting the ESP hypothesis.

RESULTS

The results are shown in Table 1. The trials scored against the unshifted E-cards were at the level of 20.75 percent hits, while those

scored against the shifted E-cards were at the level of 28.26 percent hits.

Table 1

CHI-SQUARE TEST OF THE DIFFERENCE IN SCORING RATE
BY SUBJECT P.M. IN TWO CATEGORIES OF E-CARD TRIALS

	Hits	Misses	Total	% hits
E-cards unshifted ^a	50	191	241	20.75
E-cards shifted ^b	384	975	1359	28.26
Total	434	1166	1600	

$\chi^2 = 5.47$, corrected for continuity
 $p < .01$, one-tailed

^aE-cards that the subject placed on the same pegs as they had occupied in the just-completed runs.

^bE-cards that the subject placed in positions other than where they had been during the just-completed runs.

The chi-square test of the difference between these two categories of E-card trials yields a value of 5.47 (corrected), with $p < .01$. Thus the result predicted on the ESP hypothesis was found at a level that is statistically significant. At the same time, the difference is hard to reconcile with the CM hypothesis, so the claim of the critics that the overall high rate of scoring on E-card trials was brought about through misplacement of target cards by the experimenter is not tenable as an interpretation of the Hansel effect.

The foregoing analysis completes the essential purpose of this report, and the paper could end here. However, there may be other analyses from which some further insight might be gained into the mechanism through which the high rate of scoring on the shifted E-cards was produced.

First, it is convenient to fix in mind the fact that the P.M. data forming the background of this study consist of 154 runs. Since there are two E-cards to be considered for each run, Table 1 is thus concerned with the E-piles laid down to match 308 E-cards. Of this number, 49 were unshifted and 259, shifted.

It is reasonable to assume that the unshifted E-cards were not the only ones that the subject did not remember, but that others that were not remembered were distributed in approximately the same

proportions over the other four "shifted" positions. We may estimate, therefore, that of the 259 shifted E-cards upward of 196 were not remembered and were placed in their new positions without special attention by the subject, and that only the remainder, approximately 63, were remembered and shifted intentionally. (I feel that these figures probably exaggerate the number of "forgotten" E-cards that were shifted, but the important point to be made is that shifted E-cards were made up of a population of ESP symbols the majority of which were not favorable to high ESP scoring and a smaller number that took on a special salience favoring a high rate of success. Before presenting an analysis bearing upon this point, however, I will deal with another question.)

The question arises: Was there something about having a key card on an end peg that inhibited ESP performance? If so, the low rate of scoring on the unshifted E-cards (those replaced on the same end pegs as they had been on in the preceding runs) might be due to their being in the end positions rather than to their being unshifted. To answer this question I compared the rate of success on the unshifted E-cards with that on the E-cards that were shifted to the opposite end of the row of pegs. There were 49 unshifted E-cards, and Table 1 shows that the trials made against them reached a scoring level of 20.75%. On the other hand, there were 66 E-cards that were shifted to the extreme opposite position, and they were associated with scores at the level of 32.82%, which is higher than the 28.23% rate for the shifted E-cards in general. This result shows that the low rate of success on the unshifted E-cards is not due to the end positions as such but rather to the fact that the subject failed to move them, presumably because she forgot where they were in the preceding runs.

As pointed out in the second paragraph above, not all of the E-cards that were shifted were replaced on the pegs in different positions because the subject remembered them and therefore deliberately changed their locations. If we had any way of knowing which were these "forgotten" E-cards that were shifted, they could have been eliminated from the analysis in Table 1 and this should have increased the difference in scoring rate in the two groups of E-pile trials, and the value of chi square probably would have been higher in spite of the smaller size of the sample. However, the analysis has turned out to be effective for demonstrating a clear difference between the trials made for unshifted and shifted E-cards even though the latter group is heavily contaminated with data that properly belong in the former group.

Table 2
DISTRIBUTION OF SCORING LEVELS IN E-PILES
ASSOCIATED WITH UNSHIFTED AND SHIFTED E-CARDS
(NUMBER OF E-PILES AT EACH PERCENTAGE LEVEL)

Percentage Interval	E-cards	
	Unshifted	Shifted
0- 4	17	42
5- 9	0	0
10-14	2	6
15-19	3	26
20-24	9	48
25-29	5	33
30-34	3	20
35-39	0	5
40-44	5	32
45-49	0	0
50-54	2	24
55-59	0	6
60-64	2	9
65-69	1	6
70-74	0	1
75-79	0	0
80-84	0	0
85-89	0	0
90-94	0	0
95-100	0	1
Total	49	259

In order to see whether this point need not be left as merely a logical consideration but is one that can be demonstrated to some degree on an empirical basis, I examined the distribution of scoring trends in the two groups of trials: those E-piles associated with the unshifted E-cards and the larger number of E-piles associated with the shifted E-cards. The hits made in each E-pile were expressed as a percentage of the cards (trials) in that pile. Then the distribution of these percentages was obtained for each of the groups of E-card data shown in Table 1. These two distributions of percentage of hits in the piles for unshifted and shifted E-cards are shown in Table 2.

No statistical test of the data in Table 2 will be presented because this would appear to attach more importance to these figures than they merit. It is sufficient to point out that there does appear to be a tendency for the scores from the E-piles associated with the shifted

E-cards to show a bimodal distribution that is more pronounced than can be observed in the distribution of scores in the E-piles for the unshifted E-cards. It seems reasonable to speculate, therefore, that the Hansel effect is concentrated in a relatively small proportion of the E-pile trials: those that were made against the E-cards for which the subject remembered their positions in the preceding run and took special pains to place back on the pegs in a different position.

Finally, we may ask: Was the failure to remember the E-cards due to a lack of attentiveness to the order of all of the key cards during particular runs, or might one E-card be remembered and the other forgotten with a consequent difference in the rate of success on those two cards in the next run? To attempt to answer this question, we may examine the rate of scoring when one E-card was unshifted and the other was shifted to one of the three central positions. Of the total of 49 unshifted E-cards, 10 occurred in five runs in which both E-cards were placed back in the same positions they were in previously. There were 49 trials made for these 10 E-cards and only four hits (8.16%). In the other 39 runs in which only one E-card was shifted, the rate of scoring on the unshifted cards was 23.96% while that on the 39 shifted E-cards was 27.44%. The difference in success level within these special runs is not as large as the difference that has been observed between shifted and unshifted E-cards in general, the sample is small, and the difference is not statistically significant. Therefore no answer can be given to the question: When P.M. forgot where an E-card had been (as indicated by the fact that she did not shift it), did she also forget the other one in the same run even though its position was changed?

CONCLUDING REMARKS

The relationship between ESP and memory has become a topic of current concern in parapsychology. Roll (1966) published a theoretical article, one that was developed in relation to earlier findings in the field, which put forward the hypothesis that ESP is a form of memory. More recently, Kanthamani and Rao (1974) reported experimental results which showed that significant ESP performance occurred when the concealed target was a randomly selected word from a memorized pair that the subject recalled correctly, whereas nonsignificant ESP results were obtained when the target word was from a pair that was not recalled. The evidence from the present study is consistent with their finding, since this study seems to show that the subject scored better on those E-cards

for which the former positions were remembered. Undoubtedly, however, this possible relevance of these findings does not equal in importance the main bearing of the evidence that this research has brought to light: the support that the results give to the ESP hypothesis as the correct interpretation of the Hansel effect.

REFERENCES

- GATLIN, L. L. The Pratt-Woodruff experiment and viable explanatory hypotheses. *Journal of Parapsychology*, 1975, **39**, 228-235.
- HANSEL, C. E. M. A critical analysis of the Pratt-Woodruff experiment. *Journal of Parapsychology*, 1961, **25**, 99-113.
- KANTHAMANI, H., & RAO, H. H. A study of memory-ESP relationships using linguistic forms. *Journal of Parapsychology*, 1974, **38**, 286-300.
- MEDHURST, R. G., & SCOTT, C. A re-examination of C. E. M. Hansel's criticism of the Pratt-Woodruff experiment. *Journal of Parapsychology*, 1974, **38**, 163-184.
- PRATT, J. G. Comments on the Medhurst-Scott criticism of the Pratt-Woodruff experiment. *Journal of Parapsychology*, 1974, **38**, 185-201. (a)
- PRATT, J. G. Reply to Dr. Scott. *Journal of Parapsychology*, 1974, **38**, 207-214. (b)
- PRATT, J. G., & WOODRUFF, J. L. Size of stimulus symbols in extra-sensory perception. *Journal of Parapsychology*, 1939, **3**, 121-158.
- PRATT, J. G., & WOODRUFF, J. L. Refutation of Hansel's allegation concerning the Pratt-Woodruff series. *Journal of Parapsychology*, 1961, **25**, 114-129.
- ROLL, W. G. ESP and memory. *International Journal of Neuropsychiatry*, 1966, **2**, 505-521.
- SCOTT, C. The Pratt-Woodruff experiment: reply to Dr. Pratt's comments. *Journal of Parapsychology*, 1974, **38**, 202-206.

*Division of Parapsychology
Department of Psychiatry
University of Virginia School of Medicine
Charlottesville, VA 22901*